MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Media Arts and Sciences Section

Course 4.902: Pattern Recognition and Analysis — Assigned: 2/23/93

Problem Set No. 4 — Due: 3/2/93

READING ASSIGNMENT: Therrien, Sections 5.3-6.3

This problem set consists of four problems. Remember that you are *not* allowed to look at old problem sets, but collaboration is allowed and encouraged.

In problems that call for Matlab, any software package of your choice may be used instead.

Problem 1: (25 points): Problem 5.5 in Therrien.

Hint: To show that the $\mathbf{G}_i$'s don't change, it is sufficient to show that the $\mathbf{R}_i'$'s don't change.

Problem 2: (25 points)

Let $\mathbf{A}$ be a symmetric $n \times n$ matrix with distinct nonzero eigenvalues $\lambda_1, \ldots, \lambda_n$ such that $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the corresponding eigenvectors, and let $\mathbf{x}^{(0)}$ be an an arbitrary nonzero n-dimensional vector. Consider the recursion formula

$$\mathbf{x}^{(n)} = \frac{\mathbf{A}\mathbf{x}^{(n-1)}}{|\mathbf{A}\mathbf{x}^{(n-1)}|} \qquad \text{for } n = 1, 2, 3, \ldots,$$

where $|\cdot|$ indicates magnitude.

a) Argue that $\mathbf{x}^{(n)}$ converges to one of the eigenvectors of $\mathbf{A}$. Why do we require that $|\lambda_1| \neq |\lambda_2|$? This numerical method of approximating an eigenvector is called the *power method*. (See Howard Anton, *Elementary Linear Algebra*, 4th edition, New York: Academic Press, 1984.)

b) Consider the matrix

$$\mathbf{A}' = \mathbf{A} - \lambda_1 \mathbf{e}_1 \mathbf{e}_1{}^{\mathrm{T}}.$$

What are the eigenvectors and eigenvalues of $\mathbf{A}'$ in terms of those of $\mathbf{A}$?

c) Assume now that $|\lambda_1| > |\lambda_2| > |\lambda_3| > \cdots > |\lambda_n|$, and suppose that $\mathbf{e}_1$ has already been approximated using the power method. The corresponding eigenvalue can be approximated by the *Rayleigh Quotient*

$$\lambda_1 \approx \frac{\mathbf{e}_1^{\mathrm{T}} \mathbf{A} \mathbf{e}_1}{\mathbf{e}_1^{\mathrm{T}} \mathbf{e}_1}.$$

We now wish to approximate the remaining eigenvectors. Based on your answer to part (b), describe how the power method might be modified to result in an approximation of $\mathbf{e}_2$ and λ_2. Extend your idea to cover all of the eigenvectors and eigenvalues.

d) Use the power method in Matlab to find the dominant eigenvector (the one with the largest eigenvalue) of

$$\mathbf{A} = \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix},$$

starting with $\mathbf{x}^{(0)} = [1\ 2\ 3]^{\mathrm{T}}$. Approximate the remaining eigenvectors using the procedure you developed in part (c). Compare your results with what Matlab's *eig* command gives you, and comment on any differences.

Problem 3: (25 points)

This problem explores the approach to obtaining classification features described Section 5.4.1. Suppose that the means and covariances for two classes of observation vectors are

$$\mathbf{m}_1 = [3\ \ 4]^{\mathrm{T}} \qquad \mathbf{m}_2 = [5\ \ 1]^{\mathrm{T}} \qquad \mathbf{K}_1 = \mathbf{K}_2 = \mathbf{K} = \begin{bmatrix} 1 & 1/3 \\ 1/3 & 1 \end{bmatrix}$$

a) Sketch the one-sigma ellipses in the $\mathbf{x}$ plane. (By hand would be fine, or feel free to use Matlab.)

b) What are the correlation matrices $\mathbf{R}_1$ and $\mathbf{R}_2$? On the same plot you made in part (a), indicate the directions of the dominant eigenvector for each of the matrices $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{K}$. By inspection, which of the three directions would be the best for classification?

c) Find the matrices $\mathbf{R}'_1$ and $\mathbf{R}'_2$ as defined by Eq. (5.24) in the text. Find the common eigenvectors of these two matrices, and the associated eigenvalues for each class. Does the relationship given by Eq. (5.26) hold?

d) Determine the means and covariances of the two classes in the $\mathbf{x}'$ coordinate system as defined by Eq. (5.23), and sketch the one-sigma ellipses in the $\mathbf{x}'$ plane. On your sketch, indicate the directions of the eigenvectors of $\mathbf{R}'_1$. Which do you think would be better for classification: using one of these directions in the $\mathbf{x}'$ plane, or using the direction you chose in part (b) in the $\mathbf{x}$ plane?

Problem 4: (25 points)

Consider the matrix,

$$\mathbf{A} = \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$$

a) Find the SVD of $\mathbf{A}$ in the form $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{T}}$. The SVD is not generally unique, but you should find the unique SVD that has the minimum number of minus signs in $\mathbf{U}$ and $\mathbf{V}$.

b) List the singular values σ_1, σ_2, left singular vectors $\mathbf{u}_1, \mathbf{u}_2$ (columns of $\mathbf{U}$) and right singular vectors $\mathbf{v}_1, \mathbf{v}_2$ (columns of $\mathbf{V}$). Draw a carefully labeled picture of the unit circle in $\mathbf{R}^2$ showing the vectors $\mathbf{v}_1, \mathbf{v}_2$. Draw another carefully labeled picture which shows the ellipsoid obtained from applying the mapping $\mathbf{A}$ to the unit circle, and indicate where $\mathbf{v}_1, \mathbf{v}_2$ are mapped to.

c) Find $\mathbf{A}^{-1}$ using the SVD.

d) Verify the relationship between the eigenvalues λ_1, λ_2 of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ and the singular values of $\mathbf{A}$.

e) Now consider a non-square matrix $\mathbf{A} \in \mathbf{R}^{\mathbf{m}\times\mathbf{n}}$ with $m > n$. Use the SVD to show that there exists $\mathbf{Q} \in \mathbf{R}^{\mathbf{m}\times\mathbf{n}}$ and $\mathbf{P} \in \mathbf{R}^{\mathbf{n}\times\mathbf{n}}$ such that $\mathbf{A} = \mathbf{QP}$, $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$, and $\mathbf{P}$ is symmetric positive semidefinite. This decomposition is sometimes referred to as the *polar decomposition*. It is analogous to the complex number factorization $z = e^{j\,\mathbf{arg}(z)}|z|$ where $j = \sqrt{-1}$.